

RAW SEQUENCE LISTING DATE: 02/12/2001
 PATENT APPLICATION: US/09/242,657B TIME: 10:52:56

Input Set : A:\Pto.amc
 Output Set: N:\CRF3\02122001\I242657B.raw

SEQUENCE LISTING

3 (1) GENERAL INFORMATION:

4 (i) APPLICANT: Peter Ruhdal Jensen
 5 Karin Hammer
 7 (ii) TITLE OF INVENTION: Artificial promoter libraries
 8 for selected organisms and promoters derived from such
 9 libraries
 11 (iii) NUMBER OF SEQUENCES: 58
 13 (iv) CORRESPONDENCE ADDRESS:
 14 (A) ADDRESSEE: Stanislaus Aksman
 15 Hunton & Williams
 16 (B) STREET: 1900 K Street, NW
 17 (C) CITY: Washington, DC
 18 (E) COUNTRY: USA
 19 (F) ZIP: 20006-1109
 21 (v) COMPUTER READABLE FORM:
 22 (A) MEDIUM TYPE: Floppy disk
 23 (B) COMPUTER: IBM PC compatible
 24 (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 25 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
 26 (EPO)
 28 (vi) CURRENT APPLICATION DATA:
 C--> 29 (A) APPLICATION NUMBER: US/09/242,657B
 C--> 30 (B) FILING DATE: 19-Feb-1999
 32 (vii) PRIOR APPLICATION DATA:
 33 (A) APPLICATION NUMBER: DK 886/96
 34 (B) FILING DATE: 23-AUG-1996
 35 (A) APPLICATION NUMBER: PCT/DK97/00342
 36 (B) FILING DATE: August 25, 1997
 38 (viii) ATTORNEY/AGENT INFORMATION:
 C--> 39 (C) REFERENCE/DOCKET NUMBER: 55411.000002
 41 (ix) TELECOMMUNICATION INFORMATION:
 42 (A) TELEPHONE: (202) 955-1500
 43 (B) TELEFAX: (202) 778-2201

45 (2) INFORMATION FOR SEQ ID NO: 1:

47 (i) SEQUENCE CHARACTERISTICS:
 48 (A) LENGTH: 100 base pairs
 49 (B) TYPE: nucleic acid
 50 (C) STRANDEDNESS: double
 51 (D) TOPOLOGY: linear
 53 (ii) MOLECULE TYPE: DNA (genomic)
 55 (iii) HYPOTHETICAL: YES
 56 (iv) ANTI-SENSE: NO
 58 (vi) ORIGINAL SOURCE:
 59 (A) ORGANISM: Lactococcus lactis
 61 (ix) FEATURE:
 62 (A) NAME/KEY: promoter

RAW SEQUENCE LISTING DATE: 02/12/2001
 PATENT APPLICATION: US/09/242,657B TIME: 10:52:56

Input Set : A:\Pto.amc
 Output Set: N:\CRF3\02122001\I242657B.raw

```

63      (B) LOCATION:26..82
64      (C) IDENTIFICATION METHOD: experimental
65      (D) OTHER INFORMATION:/evidence= EXPERIMENTAL
66 /standard_name= "Artificial promoter library" /note= "A
67 degenerated sequence specifying a mixture of artificial
68 promoters covering a wide range of expression in small steps
69 in L. lactis"
71      (ix) FEATURE:
72      (A) NAME/KEY: misc_feature
73      (B) LOCATION:31..45
74      (D) OTHER INFORMATION:/standard_name= "Consensus
75 sequence"
77      (ix) FEATURE:
78      (A) NAME/KEY: misc_feature
79      (B) LOCATION:60..69
80      (D) OTHER INFORMATION:/standard_name= "Consensus
81 sequence"
83      (ix) FEATURE:
84      (A) NAME/KEY: misc_feature
85      (B) LOCATION:74..82
86      (D) OTHER INFORMATION:/standard_name= "Consensus
87 sequence"
89      (ix) FEATURE:
90      (A) NAME/KEY: -35_signal
91      (B) LOCATION:40..45
92      (D) OTHER INFORMATION:/standard_name= "-35 box"
94      (ix) FEATURE:
95      (A) NAME/KEY: -10_signal
96      (B) LOCATION:63..68
97      (D) OTHER INFORMATION:/standard_name= "Pribnow box"
99      (ix) FEATURE:
100     (A) NAME/KEY: misc_recomb
101     (B) LOCATION:3..25
102     (C) IDENTIFICATION METHOD: experimental
103     (D) OTHER INFORMATION:/evidence= EXPERIMENTAL
104 /standard_name= "Multiple cloning site" /label= MCS
105 /note= "A sequence specifying recognition sites
106 for the restriction endonucleases: NlaIV, BstYI, BamHI, AlwI,
107 MboI,
108 DpnI, AflII, MseI, SspI, NsiI."
110     (ix) FEATURE:
111     (A) NAME/KEY: misc_recomb
112     (B) LOCATION:74..98
113     (C) IDENTIFICATION METHOD: experimental
114     (D) OTHER INFORMATION:/evidence= EXPERIMENTAL
115 /standard_name= "Multiple cloning site"
116 /label= MCS
117 /note= "A sequence specifying recognition sites
118 for the restriction endonucleases: ScaI, RsaI, HpaI, HincII,
```

RAW SEQUENCE LISTING DATE: 02/12/2001
 PATENT APPLICATION: US/09/242,657B TIME: 10:52:56

Input Set : A:\Pto.amc
 Output Set: N:\CRF3\02122001\I242657B.raw

```

119 MseI, SfcI,
120 PstI, Fnu4HI, BbvI, PvuII, NspBII, AluI, EcoRI."
122 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
124 CGGGATCCTT AAGAATATTA TGCATNNNNN AGTTTATTCT TGACANNNNN NNNNNNNNNT 60
126 GGTATAATAN NANAGTACTG TTAACGTCAG CTGAATTCGG 100
128 (2) INFORMATION FOR SEQ ID NO: 2:
130 (i) SEQUENCE CHARACTERISTICS:
131 (A) LENGTH: 113 base pairs
132 (B) TYPE: nucleic acid
133 (C) STRANDEDNESS: double
134 (D) TOPOLOGY: linear
136 (ii) MOLECULE TYPE: DNA (genomic)
138 (iii) HYPOTHETICAL: YES
140 (iv) ANTI-SENSE: NO
142 (ix) FEATURE:
143 (A) NAME/KEY: promoter
144 (B) LOCATION:23..95
145 (D) OTHER INFORMATION:/standard_name=
146 "Artificial promoter library"
147 /note= "A degenerated sequence specifying a mixture
148 of artificial temperature regulated promoters covering a wide
149 range of expression in small steps in L. lactis"
151 (ix) FEATURE:
152 (A) NAME/KEY: misc_feature
153 (B) LOCATION:23..49
154 (D) OTHER INFORMATION:/standard_name=
155 "Sequence providing temperature regulation to promoters"
156 /note= "This sequence comprising two inverted
157 repeats separated by a short spacer provides temperature (heat
158 shock) regulation to promoters in Gram-positive bacteria"
160 (ix) FEATURE:
161 (A) NAME/KEY: misc_feature
162 (B) LOCATION:50..60
163 (D) OTHER INFORMATION:/standard_name=
164 "Consensus sequence"
166 (ix) FEATURE:
167 (A) NAME/KEY: misc_feature
168 (B) LOCATION:75..84
169 (D) OTHER INFORMATION:/standard_name= "Consensus
170 sequence"
172 (ix) FEATURE:
173 (A) NAME/KEY: misc_feature
174 (B) LOCATION:89..95
175 (D) OTHER INFORMATION:/standard_name= "Consensus
176 sequence"
178 (ix) FEATURE:
179 (A) NAME/KEY: -35_signal
180 (B) LOCATION:55..60
181 (D) OTHER INFORMATION:/standard_name= "-35 box"

```

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/09/242,657B

DATE: 02/12/2001
 TIME: 10:52:56

Input Set : A:\Pto.amc
 Output Set: N:\CRF3\02122001\I242657B.raw

TECH CENTER 1600/269000

FEB 14 2001

RECEIVED

```

183      (ix) FEATURE:
184          (A) NAME/KEY: -10_signal
185          (B) LOCATION:78..83
186          (D) OTHER INFORMATION:/standard_name= "Pribnow box"
188      (ix) FEATURE:
189          (A) NAME/KEY: misc_recomb
190          (B) LOCATION:3..22
191          (D) OTHER INFORMATION:/standard_name= "Multiple
192 cloning site"
193 /label= MCS
194 /note= "A sequence specifying recognition sites
195 for the restriction endonucleases: NlaIV, BstYI, BamHI, AlwI,
196 MboI,DpnI, HindIII, AluI, MseI (2 sites), SspI, AseI."
198      (ix) FEATURE:
199          (A) NAME/KEY: misc_recomb
200          (B) LOCATION:89..111
201          (D) OTHER INFORMATION:/standard_name= "Multiple
202 cloning site"
203 /label= MCS
204 /note= "A sequence specifying recognition sites
205 for the restriction endonucleases: ScaI, RsaI, SfcI, PstI,
206 Fnu4HI, BbvI,
207 PvuII, NspBII, AluI, XbaI, MaeI, EcoRI, ApoI."
209      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
211 CGGGATCCAA GCTTAATATT AATTAGCACT CNNNNNNNNN GAGTGCTAAT TTTTGTGACA 60
213 NNNNNNNNNN NNNNTGGTAT AATANNANAG TACTGCAGCT GTCTAGAATT CGG      113
215 (2) INFORMATION FOR SEQ ID NO: 3:
217      (i) SEQUENCE CHARACTERISTICS:
218          (A) LENGTH: 199 base pairs
219          (B) TYPE: nucleic acid
220          (C) STRANDEDNESS: double
221          (D) TOPOLOGY: linear
223      (ii) MOLECULE TYPE: DNA (genomic)
225      (iii) HYPOTHETICAL: YES
227      (iv) ANTI-SENSE: NO
229      (vi) ORIGINAL SOURCE:
230          (A) ORGANISM: Saccharomyces cerevisiae
232      (ix) FEATURE:
233          (A) NAME/KEY: protein_bind
234          (B) LOCATION:10..16
235          (D) OTHER INFORMATION:/function= "Activating
236 promoters in
237 S. cerevisiae"
238 /bound_moiety= "GCN4 protein"
239 /standard_name= "Upstream activating sequence" /label=
240 UAS_GC4p
241 /note= "A DNA sequence that specifies a binding
242 site for the GCN4 protein, which activates the transcription
243 of genes involved in amino acid synthesis in S. cerevisiae."

```

RAW SEQUENCE LISTING DATE: 02/12/2001
 PATENT APPLICATION: US/09/242,657B TIME: 10:52:56

Input Set : A:\Pto.amc
 Output Set: N:\CRF3\02122001\I242657B.raw

```

245     (ix) FEATURE:
246         (A) NAME/KEY: TATA_signal
247         (B) LOCATION:67..72
248         (D) OTHER INFORMATION:/standard_name= "TATA box"
250     (ix) FEATURE:
251         (A) NAME/KEY: misc_signal
252         (B) LOCATION:122..144
253         (D) OTHER INFORMATION:/function= "Transcription
254 initiation"
255 /standard_name= "TI box"
257     (ix) FEATURE:
258         (A) NAME/KEY: protein_bind
259         (B) LOCATION:122..144
260         (D) OTHER INFORMATION:/bound_moiety= "Arginine
261 repressor"
262 /standard_name= "arginine repressor binding
263 site"
264 /label= argR
266     (ix) FEATURE:
267         (A) NAME/KEY: misc_RNA
268         (B) LOCATION:145..192
269         (D) OTHER INFORMATION:/function= "Spacer"
270 /standard_name= "Part of native sequence for
271 ARG8
272 gene incl. first codon"
274     (ix) FEATURE:
275         (A) NAME/KEY: misc_recomb
276         (B) LOCATION:3..8
277         (D) OTHER INFORMATION:/standard_name= "Recognition
278 site for restriction endonuclease EcoRI"
279 /label= EcoRI_site
281     (ix) FEATURE:
282         (A) NAME/KEY: misc_recomb
283         (B) LOCATION:192..197
284         (D) OTHER INFORMATION:/standard_name= "Recognition
285 site or restriction endonuclease BamHI"
286 /label= BamHI_site
288     (ix) FEATURE:
289         (A) NAME/KEY: promoter
290         (B) LOCATION:10..192
291         (D) OTHER INFORMATION:/standard_name= "Artificial
292 promoter library"
293 /note= "A degenerated sequence specifying a mixture
294 of
295 artificial promoters covering a wide range of
296 expression in small steps in S. cerevisiae"
298     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
300 CAGAATTCGT GACTCANNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN 60
302 NNNNNNNNNN NNNNNNTATA AANNNNNNNN NNNNNNNNNN NNNNNNNNNN 120

```

VERIFICATION SUMMARY

DATE: 02/12/2001

PATENT APPLICATION: US/09/242,657B

TIME: 10:52:57

Input Set : A:\Pto.amc

Output Set: N:\CRF3\02122001\I242657B.raw

L:29 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]

L:30 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]

L:39 M:220 C: Keyword misspelled or invalid format, [(C) REFERENCE/DOCKET NUMBER:]

1